Rechnernutzung in der Physik

Karlsruher Institut für Technologie
Wintersemester 2020/21

Teil – 1 “Computerbasierte Methoden der Statistik”

Roger Wolf (roger.wolf@cern.ch), Ralf Ulrich (ralf.ulrich@kit.edu)

Priv. Doz. Dr. Roger Wolf
http://ekpwww.physik.uni-karlsruhe.de/~rwolf/

Inhalt

Roger Wolf

Ralf Ulrich

Literaturempfehlungen

- **Einführende Literatur zu Statistik und Numerik:**
  - G. Cowan, *Statistical data analysis*, Oxford (1997) (KIT-Bibliothek).
  - G. Bohm, G. Zech, *Einführung in Statistik und Messwertanalyse für Physiker*, DESY (2006) (eBook deutsch, eBook english).
  - V. Blobel, E. Lormann, *Statistische und numerische Methoden der Datenanalyse*, DESY (2012) (Webseite).
  - R. J. Barlow, *Statistics: A Guide to the use of statistical methods in the physical sciences*, Wiley (1989) (KIT-Bibliothek).
  - W. H. Press, S. A. Teukolsky, W. T. Vetterling, B. P. Flannery, *Numerical recipes*, Cambridge Univ. Press (2007) (Webseite).

- Skriptensammlung von Prof. G. Quast (Link):

- S. Weinzierl, *Introduction to Monte Carlo methods* (arxiv:hep-ph/0006269):

2 Monte Carlo Methode als numerisches Hilfsmittel

2.1 Einführung und Zufallszahlengeneratoren

Der Beginn jeder Anwendung der Monte Carlo Methode ist eine Kette gleichverteilter Zufallszahlen. Wie auf dem Rouletttisch sind Gleichverteilung und Unabhängigkeit der Zufallszahlen in der Sequenz eine essentielle Voraussetzung.

Die Monte Carlo (MC) Methode

- Numerisches Verfahren aus der Stochastik um analytisch nicht oder nur aufwändig lösbare Probleme mit Hilfe stochastischer Konzepte zu lösen.

- Anwendungsgebiete:
  - Numerische Mathematik (Integration, Optimierung, Faltung, …).
  - Angewandte Statistik (Bestimmung von Korrelationen, Fehlerfortpflanzung, Hypothesentests, …).
  - Nachbildung komplexer Prozesse mit statistischem Verhalten (Vielteilchensysteme, Teilchenphysik, …).

- Historie:
  - Erste Idee: Enrico Fermi 1930er Jahre.
  - Erste Ausführung: Nicolas Metropolis, Stanislaw Ulam, John von Neumann 1940er Jahre (im Rahmen des Los Alamos Projekts).
  - Namensgebung durch John von Neumann (als code Name innerhalb des Projekts).

(Pseudo-)Zufallszahlen

- Die Monte Carlo Methode basiert auf dem Gesetz der großen Zahlen. Sie beginnt mit einer Reihe gleichverteilter Zufallszahlen.

- Diese Reihe kann entweder physikalisch oder mit Hilfe eines Zufallszahlengenerators als **Pseudo-Zufallszahlen** bestimmt werden.

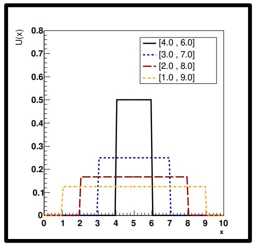
(Pseudo-)Zufallszahlen:

sind

Ergebnis einer deterministischen Sequenz (insb. reproduzierbar)

sind

Innerhalb eines vorgegebenen Intervalls nach bester Möglichkeit **gleichverteilt**.



- In einem zweiten Schritt werden die gleichverteilten (Pseudo-)Zufallszahlen in eine beliebige Wahrscheinlichkeitsdichte transformiert.

Zufallszahlengeneratoren

- **Beispiel**: multiplikative Kongruenzgeneratoren

$$y_i = \left(\left(\sum_{k=1}^{n} a_k y_{i-k} \right) + b \right) \mod m$$

$$\begin{array}{ll}
\text{Für } i > n\,, \quad n \in \mathbb{N}^+, \quad m \in \{2,3,4,5,\ldots\} & \text{(Modul)} \\
a_1, a_2, \ldots a_n \in \{0,1,2,\ldots m-1\}\,,\ a_n > 0 & \text{(Faktoren)} \\
b \in \{0,1,2,\ldots m-1\} & \text{(Inkrement)} \\
y_1, y_2, \ldots y_n \in \{0,1,2,\ldots m-1\} & \text{(Startwerte)}^{(*)}
\end{array}$$

(*) Dürfen nicht alle 0 sein wenn b=0.

- Zustand des Generators vor Erzeugung des Wertes y_i bei Vorgabe von $\left(\{a_k\}, b, m, n\right)$ durch Werte $y_{i-n}, \ldots y_{n-1}$ vorgegeben. Startwerte (→ engl. *seeds*).

- Wichtigste wünschenswerte Eigenschaften:
  - Lange Periode, bevor sich die Sequenz der Pseudo-Zufallszahlen wiederholt.
  - Möglichst keine Korrelation zwischen (unmittelbar) aufeinanderfolgenden Werten.

Linearer Kongruenzgenerator (LCG)

- Einfacher aber sehr effizienter Algorithmus aus dieser Gruppe:

$$\boxed{y_{i+1} = a \cdot y_i \mod m}$$

Zuerst eingeführt von Derrick Henry Lemmer, 1949.

- Bsp.: Durch Multiplikation mit a wird aus y_i eine neue Zahl erzeugt, danach werden (für gerades m) alle signifikanten Stellen (**A**) bis auf $\log_2(m)$ (**B**) abgeschnitten.

- **Beispiel**: $a = 12345678, \; m = 98765432$

seed index	value
1	1
2	12345678
3	61728396
4	86419752
5	12345680
6	86419752
7	12345680
8	86419752
9	...

Ein lauffähiges RooT macro finden Sie hier.

- **Problem**: Probabalistische Eigenschaften hängen von Wahl von (a, m) ab und können sehr ungünstig ausfallen.

- Mögliche Abhilfe:
  - Wähle m als möglichst große Primzahl und a als Primitivwurzel aus m .

$$\underbrace{\texttt{1011010101010}}_{\textbf{A}}\underbrace{\texttt{11101010101010001010101 0}}_{\textbf{B}}$$

Linearer Kongruenzgenerator (LCG)

- **Beispiel**:

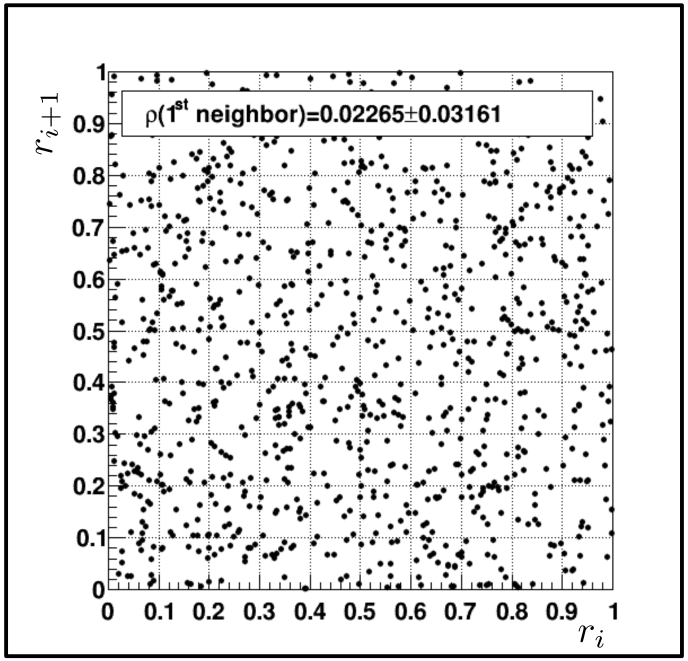
```
long int mlc_random(long int seed, int scale=40692, int modus=2147483399){
  return (scale*seed)%modus;
}
```

Ein lauffähiges RooT macro finden Sie hier.

$$\boxed{y_{i+1} = a \cdot y_i \mod m}$$

$$m = 2147483399 = 2^{31} - 249$$
$$a = 40692$$
$$y_i \to r_i = \frac{y_i}{m}$$

- Korrelation aufeinanderfolgender Werte: (*)



Priv. Doz. Dr. Roger Wolf
http://ekpwww.physik.uni-karlsruhe.de/~rwolf/

(*)
Aus jewels 1000 iteration.

Linearer Kongruenzgenerator (LCG)

- **Beispiel**:

```
long int mlc_random(long int seed, int scale=40692, int modus=2147483399){
  return (scale*seed)%modus;
}
```

Ein lauffähiges RooT macro finden Sie hier.

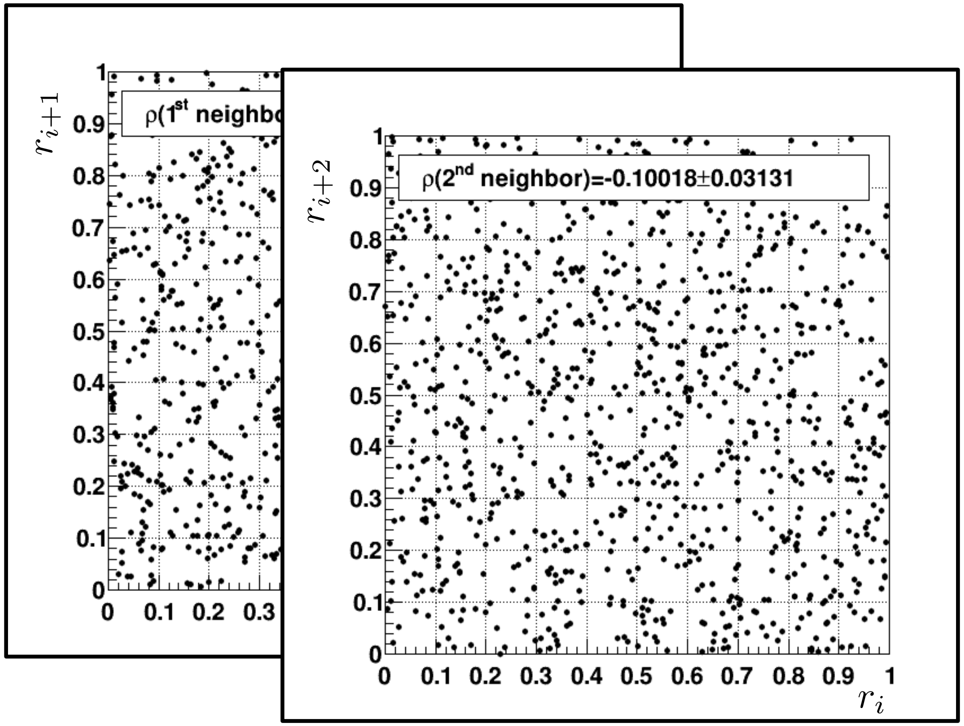
$$\boxed{y_{i+1} = a \cdot y_i \mod m}$$

$$m = 2147483399 = 2^{31} - 249$$

$$a = 40692$$

$$y_i \to r_i = \frac{y_i}{m}$$

- Korrelation aufeinanderfolgender Werte: (*)



Priv. Doz. Dr. Roger Wolf
http://ekpwww.physik.uni-karlsruhe.de/~rwolf/

(*)
Aus jewels 1000 iteration.

Linearer Kongruenzgenerator (LCG)

- **Beispiel**:

```
long int mlc_random(long int seed, int scale=40692, int modus=2147483399){
  return (scale*seed)%modus;
}
```

Ein lauffähiges RooT macro finden Sie hier.

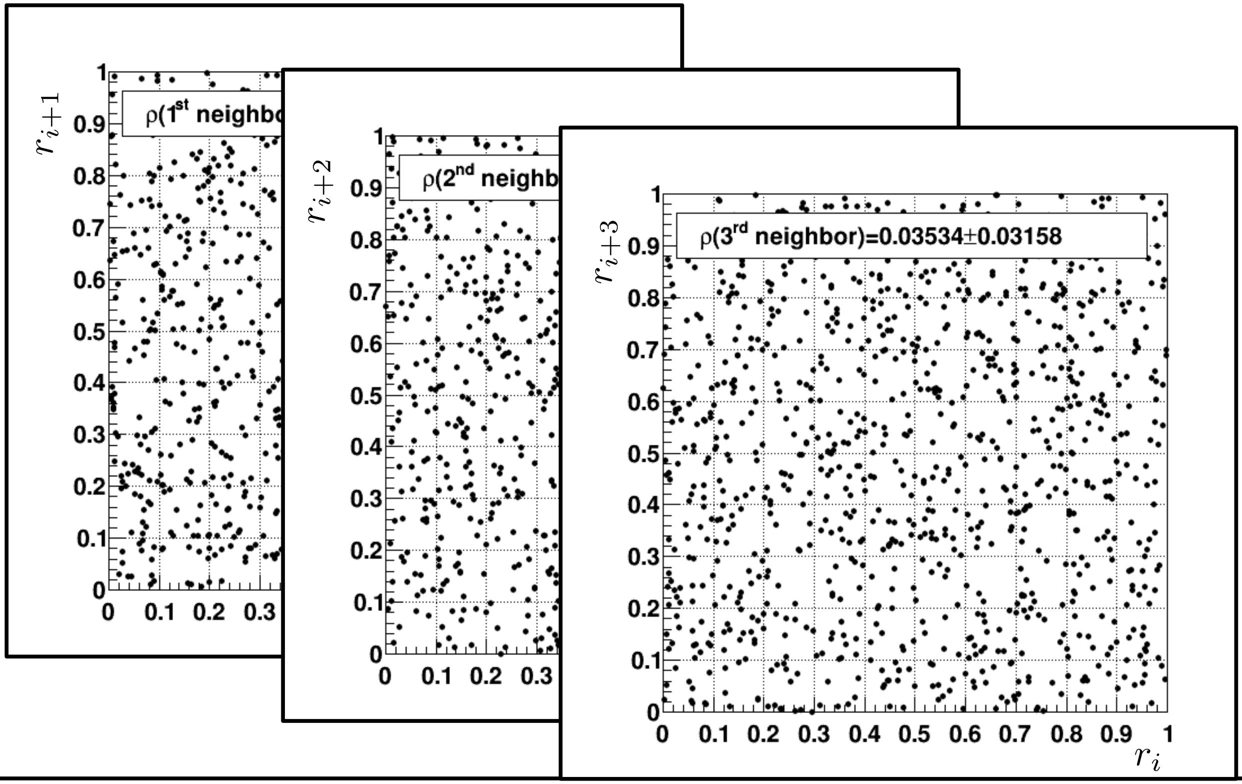
$$y_{i+1} = a \cdot y_i \mod m$$

$$m = 2147483399 = 2^{31} - 249$$

$$a = 40692$$

$$y_i \rightarrow r_i = \frac{y_i}{m}$$

- Korrelation aufeinanderfolgender Werte: (*)



Priv. Doz. Dr. Roger Wolf
http://ekpwww.physik.uni-karlsruhe.de/~rwolf/

(*)
Aus jewels 1000 iteration.

Linearer Kongruenzgenerator (LCG)

- **Beispiel**:

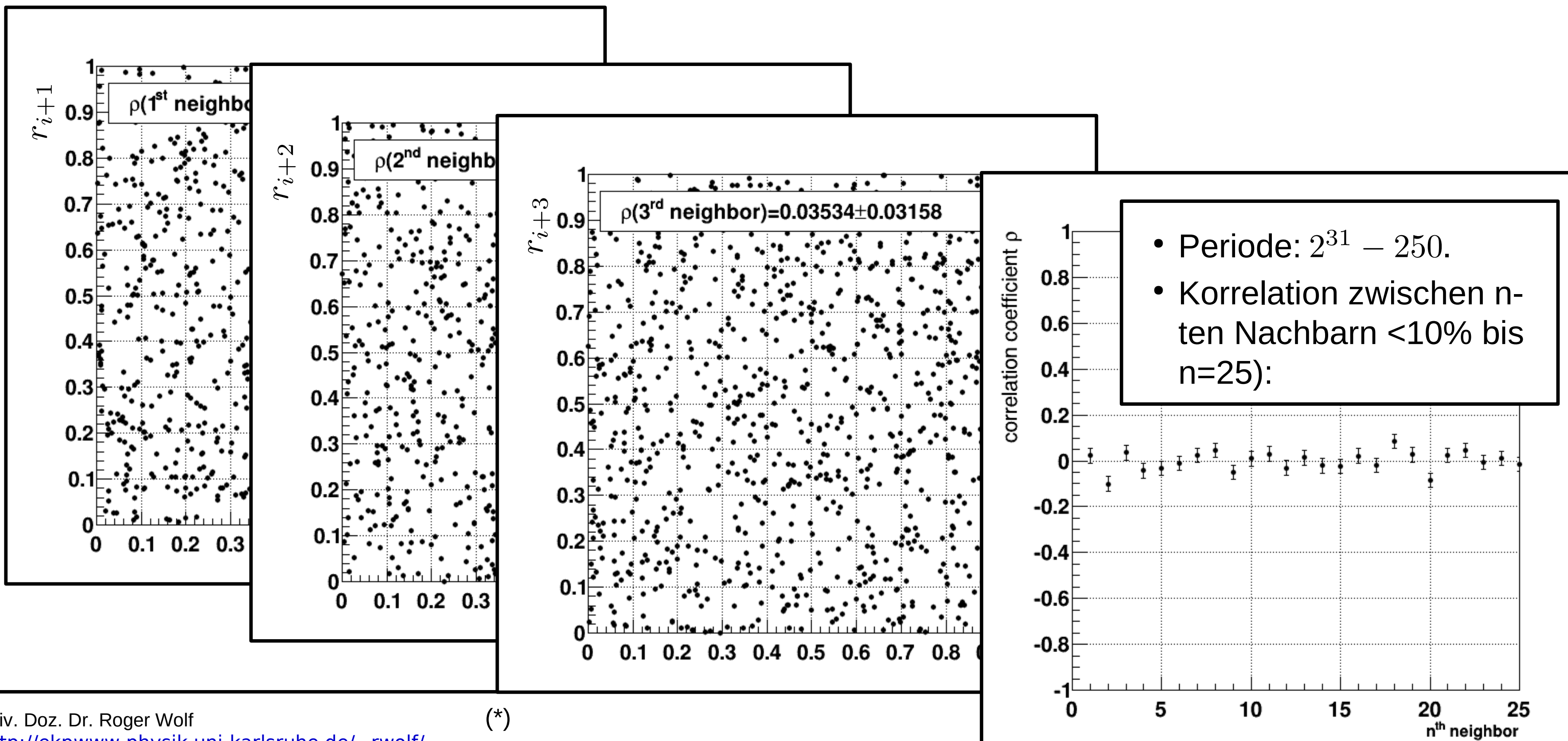
```
long int mlc_random(long int seed, int scale=40692, int modus=2147483399){
  return (scale*seed)%modus;
}
```

Ein lauffähiges RooT macro finden Sie hier.

$$y_{i+1} = a \cdot y_i \mod m$$

$$m = 2147483399 = 2^{31} - 249$$
$$a = 40692$$
$$y_i \to r_i = \frac{y_i}{m}$$

- Korrelation aufeinanderfolgender Werte: (*)



(*) Aus jewels 1000 iteration.

Moderne Zufallszahlengeneratoren

- State of the Art Bsp.-1: Mersenne-Twister.
- Basierend auf Mersenne Primzahlen: $\mathcal{M}_n = 2^n - 1$, benannt nach Marin Mersenne (1588 – 1648).
- 264 seeds (z.B. durch LCG initialisiert).
- Periode: $2^{19937} - 1$
- Je nach Initialisierung garantiert unkorreliert bis zur Dimension 263.
- Z.B. verwendet in der ROOT Klasse TRandom3.

Wikipedia

```
#include <stdint.h>

uint32_t mersenne_twister() {
#define N     624
#define M     397
#define HI    0x80000000
#define LO    0x7fffffff
  static const uint32_t A[2] = { 0, 0x9908b0df };
  static uint32_t y[N];
  static int index = N+1;
  static const uint32_t seed = 5489UL;
  uint32_t  e;

  if (index > N) {
    int i;
    /* Initialisiere y mit Pseudozufallszahlen */
    y[0] = seed;

    for (i=1; i<N; ++i) {
      y[i] = (1812433253UL * (y[i-1] ^ (y[i-1] >> 30)) + i);
      /* See Knuth TAOCP Vol2. 3rd Ed. P.106 for multiplier. */
      /* In the previous versions, MSBs of the seed affect   */
      /* only MSBs of the array mt[].                        */
      /* 2002/01/09 modified by Makoto Matsumoto             */
    }
  }

  if (index >= N) {
    int i;
    /* Berechne neuen Zustandsvektor */
    uint32_t h;
    ...
```

Moderne Zufallszahlengeneratoren

- State of the Art Bsp.-2:
  Permuted congruential generators.

- Ausgangspunkt: LCG.

- Weiterentwicklung: Nutze signifikante Bits (**A**), um nicht-signifikante Bits (**B**) zusätzlich zu permutieren.

- Erreicht verbesserte statistische Eigenschaften und nicht/kaum vorhersagbare Zahlenfolgen.

- Z.B. verwendet in numpy.random.

Wikipedia

```c
#include <stdint.h>
static uint64_t       state      = 0x4d595df4d0f33173;
static uint64_t const multiplier = 6364136223846793005u;
static uint64_t const increment  = 1442695040888963407u;

static uint32_t rotr32(uint32_t x, unsigned r)
{
    return x >> r | x << (-r & 31);
}

uint32_t pcg32(void)
{
    uint64_t x = state;
    unsigned count = (unsigned)(x >> 59);       // 59 = 6

    state = x * multiplier + increment;
    x ^= x >> 18;                               // 18 = (
    return rotr32((uint32_t)(x >> 27), count);  // 27 = 3
}

void pcg32_init(uint64_t seed)
{
    state = seed + increment;
    (void)pcg32();
}
```

$$\underbrace{1011010101010}_{\textbf{A}}\underbrace{111010101010100010101010}_{\textbf{B}}$$

Anwendung und Tests

- Neben der Anwendung im Bereich der Statistik spielen Zufallszahlengeneratoren v.a. eine Rolle für Verschlüsselungsverfahren.

- Eine wichtige weitere Eigenschaft in diesen Bereichen (die für uns keine Rolle spielt) ist die der Vorhersagbarkeit der Sequenz.

- Testsuiten werden dazu verwendet die Eigenschaften verschiedener Zufallszahlengenerato-ren zu überprüfen. Beispiele sind im folgenden verlinkt:

  - TestU01 (2007).

  - Diehard (1995).

  - Dieharder (2020).

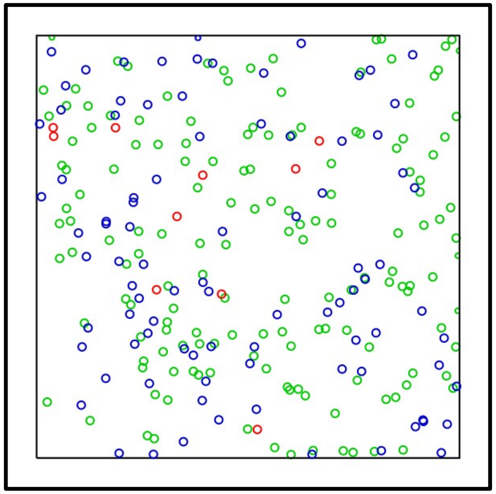
Quasizufallszahlen

- Reihen von (Pseudo-)Zufallszahlen verteilen sich statistisch in dem Raum, in dem sie bestimmt werden.

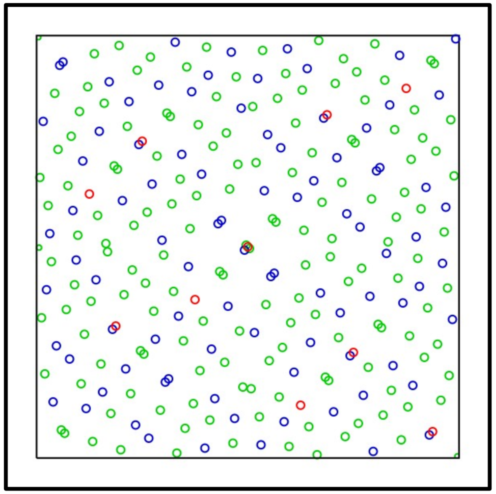
- Wenn Sie einen bestimmten Phasenraum möglichst gleichmäßig „ausleuchten“ möchten, können Sie dies effizienter mit Hilfe **korrelierter Quasizufallszahlen** erreichen.

- Beispiel: Sobol Sequenz (1964 zur effizienten Bestimmung numerischer Integrale eingeführt):

256 Pseudozufallszahlen (in 2d)



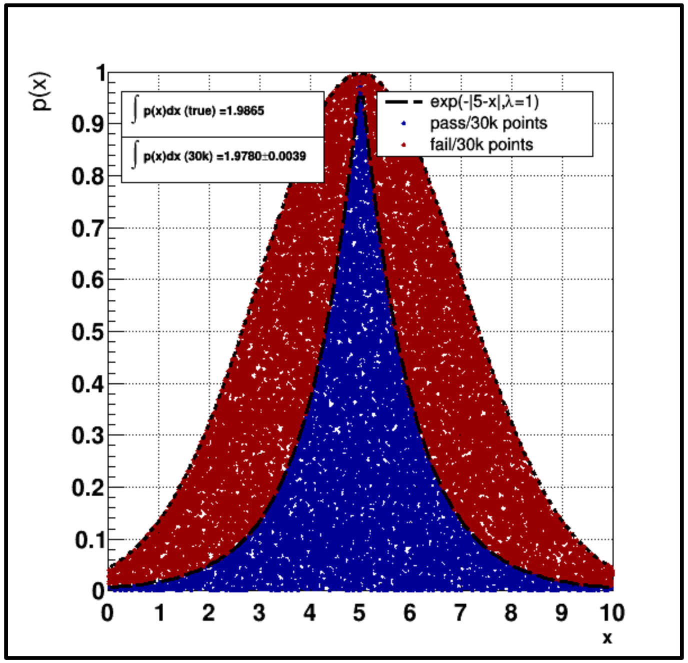
256 erste Zahlren der (2,3) Sobol Sequenz (in 2d)



- Rot: 1 … 10
- Blau: 11 … 100
- Grün: 101 … 256

2.2 Abbildung von Wahrscheinlicheitsdichten mit Hilfe gleichverteilter Zufallszahlen

Um eine beliebige Wahrscheinlichkeitsverteilung der erzeugten Zufallszahlen darstellen zu können, ist der nächste Schritt, die gleichverteilten Zufallszahlen auf eine beliebige Wahrscheinlichkeitsdichte transformieren zu können.



Zufallszahl → Wahrscheinlichkeitsdichte

- Die Wertemenge eines Zufallszahlengenerators ist nach bestem Vermögen **gleichverteilt**.

- Wie kommt man von einer Gleichverteilung von Zahlen auf eine *beliebige* Wahrscheinlichkeitsdichte?

- Zwei Methoden werden im folgenden diskutiert:
  - Die Transformationsmethode.
  - Die Verwerfungsmethode.

Zurück zu Folie 12

1. Transformationsmethode

- Finde eine Transformation $y_i = y(x_i),\ x_i \in \mathrm{U}_{[0,1]}$ nach der die y_i geeignet verteilt sind durch analytische Transformation.

- **Problemstellung**: $p(x) = \mathrm{U}_{[0,1]}(x)$ (Gleichverteilung) und $q(y)$ (gewünschte Wahrschein-lichkeitsdichte) sind vorgegeben, $y(x)$ ist gesucht.

Zurück zu Folie 12

1. Transformationsmethode

- Finde eine Transformation $y_i = y(x_i),\, x_i \in \mathrm{U}_{[0,1]}$ nach der die y_i geeignet verteilt sind durch analytische Transformation.

- **Problemstellung**: $p(x) = \mathrm{U}_{[0,1]}(x)$ (Gleichverteilung) und $q(y)$ (gewünschte Wahrscheinlichkeitsdichte) sind vorgegeben, $y(x)$ ist gesucht.

Lösung:

$$\boxed{y(x) = \mathcal{Q}^{-1}(y)}$$

1. Transformationsmethode

- Finde eine Transformation $y_i = y(x_i),\ x_i \in \mathrm{U}_{[0,1]}$ nach der die y_i geeignet verteilt sind durch analytische Transformation.

- **Problemstellung**: $p(x) = \mathrm{U}_{[0,1]}(x)$ (Gleichverteilung) und $q(y)$ (gewünschte Wahrscheinlichkeitsdichte) sind vorgegeben, $y(x)$ ist gesucht.

- **Lösungsansatz**:

$$\underbrace{\mathcal{P}(x' \leq x)}_{\text{CDF für } p(x)} = \int_0^x p(x')\,\mathrm{d}x' \qquad \underbrace{\mathcal{Q}(y(x') \leq y(x))}_{\text{CDF für } q(y(x))} = \int_{y(0)}^{y(x)} q(y(x'))\,\mathrm{d}y(x')$$

Erhaltung der Wahrscheinlichkeit:

$$\mathcal{P}(x) = \int_0^x 1\,\mathrm{d}x' = x = \int_{y(0)}^{y(x)} q(y')\,\mathrm{d}y' = \mathcal{Q}(y(x))\,; \qquad x(y) = \mathcal{Q}(y(x))\,;$$

Lösung:

$$\boxed{y(x) = \mathcal{Q}^{-1}(y)}$$

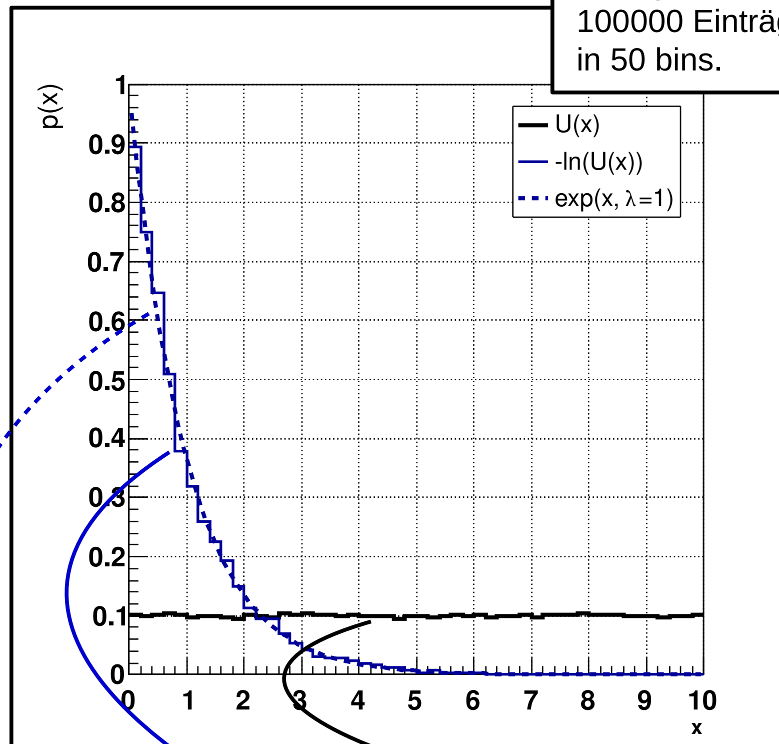
Priv. Doz. Dr. Roger Wolf
http://ekpwww.physik.uni-karlsruhe.de/~rwolf/

1. Transformationsmethode – Bsp.: 1 –

- **Beispiel-1**: $\exp(y,\lambda) = \frac{1}{\lambda}e^{-y/\lambda}$ (gewünschte Wahrscheinlichkeitsdichte für $y = 0\ldots\infty$)

Ein lauffähiges RooT macro finden Sie hier.

Histogramm mit 100000 Einträgen in 50 bins.



$\exp(x,\lambda=1)$

Befüllt mit $-\ln(x) : \mathrm{U}_{[0,1]}(x)$

Befüllt mit $\frac{1}{10}\mathrm{U}_{[0,1]}(10\,x)$

$$q(y) = \exp(y,\lambda) = \frac{1}{\lambda}e^{-y/\lambda}$$

$$x(y) = \mathcal{Q}(y(x)) = \left[-e^{-y'/\lambda}\right]_0^y = \left(1 - e^{-y/\lambda}\right)$$

$$y(x) = -\lambda\ln\left(1-x\right)$$

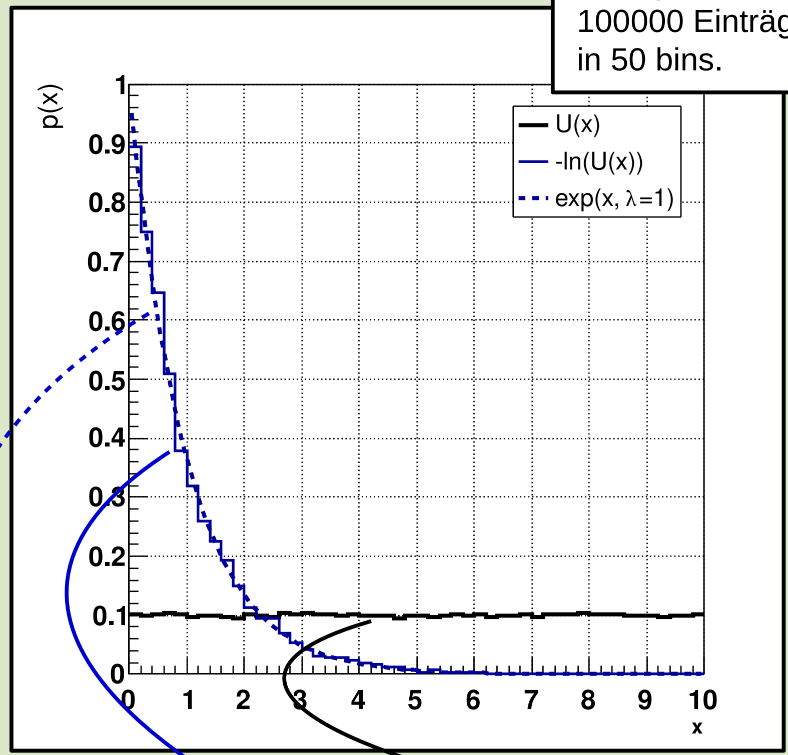
Priv. Doz. Dr. Roger Wolf
http://ekpwww.physik.uni-karlsruhe.de/~rwolf/

1. Transformationsmethode – Bsp.: 1 –

- **Beispiel-1**: $\exp(y,\lambda) = \dfrac{1}{\lambda}e^{-y/\lambda}$ (gewünschte Wahrscheinlichkeitsdichte für $y = 0\ldots\infty$)

Ein lauffähiges RooT macro finden Sie hier.

Histogramm mit 100000 Einträgen in 50 bins.



$\exp(x,\lambda=1)$

Befüllt mit $-\ln(x) : \mathrm{U}_{[0,1]}(x)$

Befüllt mit $\dfrac{1}{10}\mathrm{U}_{[0,1]}(10\,x)$

$$q(y) = \exp(y,\lambda) = \frac{1}{\lambda}e^{-y/\lambda}$$

$$x(y) = \mathcal{Q}(y(x)) = \left[-e^{-y'/\lambda}\right]_0^y = \left(1 - e^{-y/\lambda}\right)$$

$$y(x) = -\lambda\ln\left(1 - x\right)$$

NB: Die Umkehrung ist nicht eindeutig: $x(y)$ nimmt Werte zwischen 0 und 1 an. $y(x)$ ist ebenso exponentiell verteilt wie $y(1-x)$.

1. Transformationsmethode – Bsp.: 2 –

- **Beispiel-2**: $\varphi(x,\mu,\sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}$ (gewünschte Wahrscheinlichkeitsdichte)

o.B.d.A. setzen wir $\mu = 0$. Das Integral von φ lässt sich in 2d analytisch lösen:

$$\mathcal{Q}(R) = \int \frac{1}{2\pi\sigma^2} e^{-\frac{1}{2\sigma^2}\left(x_1^2+x_2^2\right)} \mathrm{d}x_1 \mathrm{d}x_2 = \frac{1}{2\pi\sigma^2} \int\limits_0^{2\pi} \mathrm{d}\phi \int\limits_0^{R} e^{-\frac{r^2}{2\sigma^2}} r\,\mathrm{d}r\,;$$

mit: $r = \sqrt{x_1^2 + x_2^2}\,;\quad \mathrm{d}r^2 = 2\,r\mathrm{d}r$

$$\mathcal{Q}(R) = \frac{1}{2\pi\sigma^2} \int\limits_0^{2\pi} \mathrm{d}\phi \int\limits_0^{R} e^{-\frac{r^2}{2\sigma^2}} r\,\mathrm{d}r = \frac{1}{\sigma^2} \int\limits_0^{R} e^{-\frac{r^2}{2\sigma^2}} \frac{\mathrm{d}r^2}{2} = \frac{1}{\sigma^2}\left[-\sigma^2 e^{-\frac{r^2}{2\sigma^2}}\right]_0^R = \left(1 - e^{-\frac{R^2}{2\sigma^2}}\right)$$

Umkehrfunktion (in 2d) für $u \in \mathrm{U}_{[0,1]}$:

$$u(R) = \left(1 - e^{-\frac{R^2}{2\sigma^2}}\right) = \mathcal{Q}(R)$$

$$R(u) = \sqrt{-2\sigma^2 \ln\left(1-u\right)} \overset{(*)}{\equiv} \sqrt{-2\ln\left(u\right)} \cdot \sigma = \mathcal{Q}^{-1}(u)$$

(*) Hier gilt das gleiche wie auf Folie 14a: mit der Transformation $u \to 1-u$ durchlaufen Sie das uniform verteilte Interall effektiv nicht von 0 bis 1 sondern von 1 bis 0.

1. Transformationsmethode – Bsp.: 2 continued –

- **Beispiel-2**: $\varphi(x,\mu,\sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}$ (gewünschte Wahrscheinlichkeitsdichte)

Wählen Sie $u \in \mathrm{U}_{[0,1]}$ und $\phi \in \mathrm{U}_{[0,2\pi]}$ gleichverteilt erhalten Sie zwei voneinander unabhängige normalverteilte Zufallszahlen:

$$x_1 = \sqrt{-2\ln(u)} \cdot \sigma \cos\phi$$
$$x_2 = \sqrt{-2\ln(u)} \cdot \sigma \sin\phi$$

Verschiebung um μ erfolgt über $x_i \to x'_i = x_i - \mu$

Warum muss ich zwei Zufallszahlen ziehen, um $\varphi(x,\mu,\sigma)$ zu samplen?

Wir haben zur Integration das Problem von einer auf zwei Dimensionen erweitert (erste Zeile der Rechnung auf Folie 15). Das entspricht dem Produkt zweier identischer Normalverteilungen in zwei unabhängigen Variablen x_1 und x_2. Danach haben wir $R(u), u \in U_{[0,1]}$ ausgerechnet. Aber φ hängt nicht von R ab, sondern von $x_1 = R\cos\phi$. Alternativ kann ich eine unabhängige ebenfalls nach φ verteilte Zufallsgröße $x_2 = R\sin\phi$ bestimmen.

1. Transformationsmethode – Bsp.: 2 continued –

- **Beispiel-2**: $\varphi(x,\mu,\sigma) = \dfrac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2} \quad \text{(gewünschte Wahrscheinlichkeitsdichte)}$

- Ein numerisch einfacheres Verfahren, bei dem Sie keine trigonometrischen Funktionen auswerten müssen, erhalten Sie über das folgende algorithmische Vorgehen:

  - Wähle zwei gleichverteilte Zufallszahlen $u_1, u_2 \in \mathrm{U}_{[0,1]}$;

  - Transformiere $v_i = 2u_i - 1$, auf diese Weise sind beide v_i gleichförmig zwischen -1 und +1 verteilt;

  - Verwerfe das Wertepaar, wenn $v^2 = v_1^2 + v_2^2 > 1$, auf diese Weise werden die v_i auf die Einheitskreisfläche um den Ursprung beschränkt;

- Sie erhalten so zwei voneinander unabhängige normalverteilte Zufallszahlen aus:

$$
\begin{aligned}
x_1 &= \sqrt{-2\frac{\ln\left(v^2\right)}{v^2}} \cdot \sigma v_1 - \mu = \sqrt{-2\ln\left(v^2\right)} \cdot \sigma \frac{v_1}{v} - \mu \\
x_2 &= \sqrt{-2\frac{\ln\left(v^2\right)}{v^2}} \cdot \sigma v_2 - \mu = \sqrt{-2\ln\left(v^2\right)} \cdot \sigma \frac{v_2}{v} - \mu
\end{aligned}
$$

1. Transformationsmethode – Bsp.: 2 continued –

- **Beispiel-2**: $\varphi(x,\mu,\sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2} \quad \text{(gewünschte Wahrscheinlichkeitsdichte)}$

- Ein numerisch einfacheres Verfahren, bei dem Sie keine trigonometrischen Funktionen auswerten müssen, erhalten Sie über das folgende algorithmische Vorgehen:

  - Wähle zwei gleichverteilte Zufallszahlen $u_1, u_2 \in \mathrm{U}_{[0,1]}$;

  - Transformiere $v_i = 2u_i - 1$, auf diese Weise sind beide v_i gleichförmig zwischen -1 und +1 verteilt;

  - Verwerfe das Wertepaar, wenn $v^2 = v_1^2 + v_2^2 > 1$, auf diese Weise werden die v_i auf die Einheitskreisfläche um den Ursprung beschränkt;

- Sie erhalten so zwei voneinander unabhängige normalverteilte Zufallszahlen aus:

$$x_1 = \sqrt{-2\frac{\ln\left(v^2\right)}{v^2}} \cdot \sigma v_1 - \mu = \sqrt{-2\ln\left(v^2\right)} \cdot \sigma\frac{v_1}{v} - \mu$$
$$x_2 = \sqrt{-2\frac{\ln\left(v^2\right)}{v^2}} \cdot \sigma v_2 - \mu = \sqrt{-2\ln\left(v^2\right)} \cdot \sigma\frac{v_2}{v} - \mu$$

Der Faktor $1/v^2$ normiert einen Kreis mit Radius $v \leq 1$ zurück auf einen Einheitskreis mit $v = 1$.

1. Transformationsmethode – Bsp.: 2 continued –

- **Beispiel-2**: $\varphi(x,\mu,\sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}$ $\quad$ (gewünschte Wahrscheinlichkeitsdichte)

- Ein numerisch einfacheres Verfahren, bei dem Sie keine trigonometrischen Funktionen auswerten müssen, erhalten Sie über das folgende algorithmische Vorgehen:
  - Wähle zwei gleichverteilte Zufallszahlen $u_1, u_2 \in \mathrm{U}_{[0,1]}$;
  - Transformiere $v_i = 2u_i - 1$, auf diese Weise sind beide v_i gleichförmig zwischen -1 und +1 verteilt;
  - Verwerfe das Wertepaar, wenn $v^2 = v_1^2 + v_2^2 > 1$, auf diese Weise werden die v_i auf die Einheitskreisfläche um den Ursprung beschränkt;

- Sie erhalten so zwei voneinander unabhängige normalverteilte Zufallszahlen aus:

$$x_1 = \sqrt{-2\frac{\ln\left(v^2\right)}{v^2}} \cdot \sigma v_1 - \mu = \sqrt{-2\ln\left(v^2\right)} \cdot \sigma \frac{v_1}{v} - \mu$$

$$x_2 = \sqrt{-2\frac{\ln\left(v^2\right)}{v^2}} \cdot \sigma v_2 - \mu = \sqrt{-2\ln\left(v^2\right)} \cdot \sigma \frac{v_2}{v} - \mu$$

$$\frac{v_2}{v} = \sin\phi$$

Der Faktor $1/v^2$ normiert einen Kreis mit Radius $v \leq 1$ zurück auf einen Einheitskreis mit $v = 1$.

1. Transformationsmethode – Diskussion –

- Die Transformationsmethode ist ohne großen Rechenaufwand leicht und algorithmisch schnell anwendbar.

- Ein Nachteil liegt in der begrenzten Anwendbarkeit, denn die Wahrscheinlichkeitsdichte muss analytisch geschlossen formulierbar und integrabel sein. Das Integral ist dann im allgemeinen bereits umkehrbar.

Zurück zu Folie 12

2. Verwerfungsmethode (engl. *stochastic sampling*)

Zuerst vorgeschlagen von John v. Neumann (1951).

- Zur Transformation in eine beliebige Wahrscheinlichkeitsdichte $p(x)$ erzeuge ein Paar gleichverteilter Zufallszahlen (x, y).

- Akzeptiere Ereignis wenn $y \leq p(x)$ und verwerfe sonst.

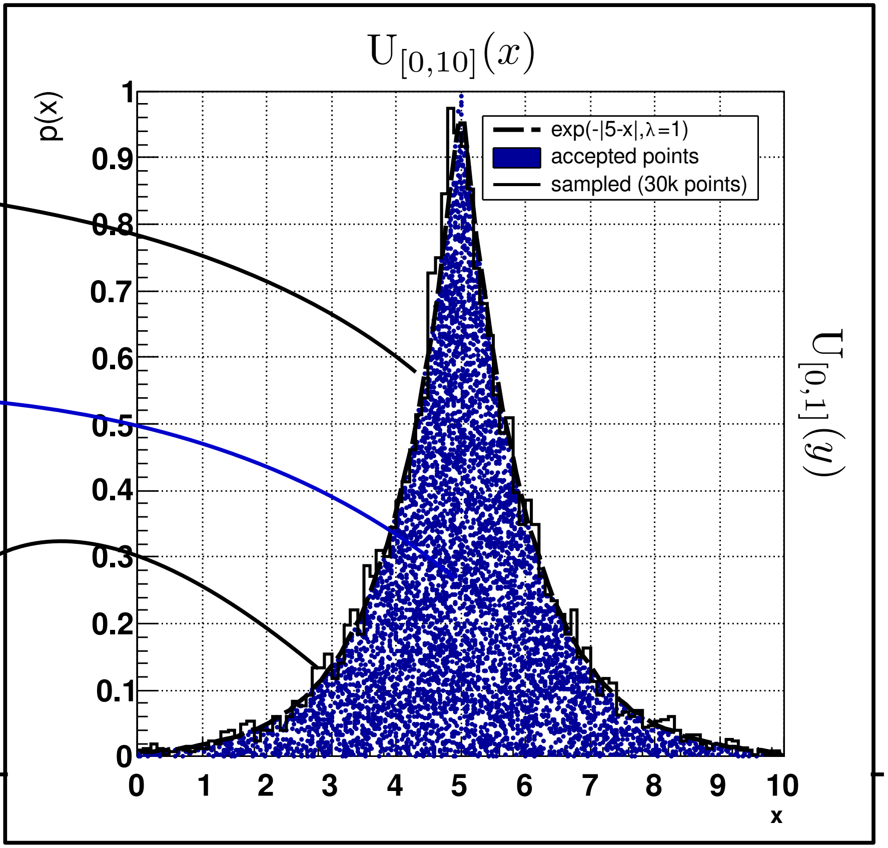
- Beispiel Funktion: $p(x) = \exp(-|5-x|)$

Ursprüngliche Wahrscheinlich-keitsdichte

Akzeptierte Inte-grationspunkte (30k stochastic sampling).

Abgeleitetes Histogramm (30k Integrationspunkte)

Ein lauffähiges RooT macro finden Sie hier.



Beispiel – Numerische Integration

- Wie läßt sich diese Methode zur numerischen Integration nutzen?
  - Teile Ereignisse ein in *pass* und *fail*.
  - Das Integral ergibt sich aus: $\int p(x)\,\mathrm{d}x = 10 \cdot \frac{n(\mathrm{pass})}{n(\mathrm{pass}) + n(\mathrm{fail})}$

In diesem Bsp. Fläche(blau+rot).

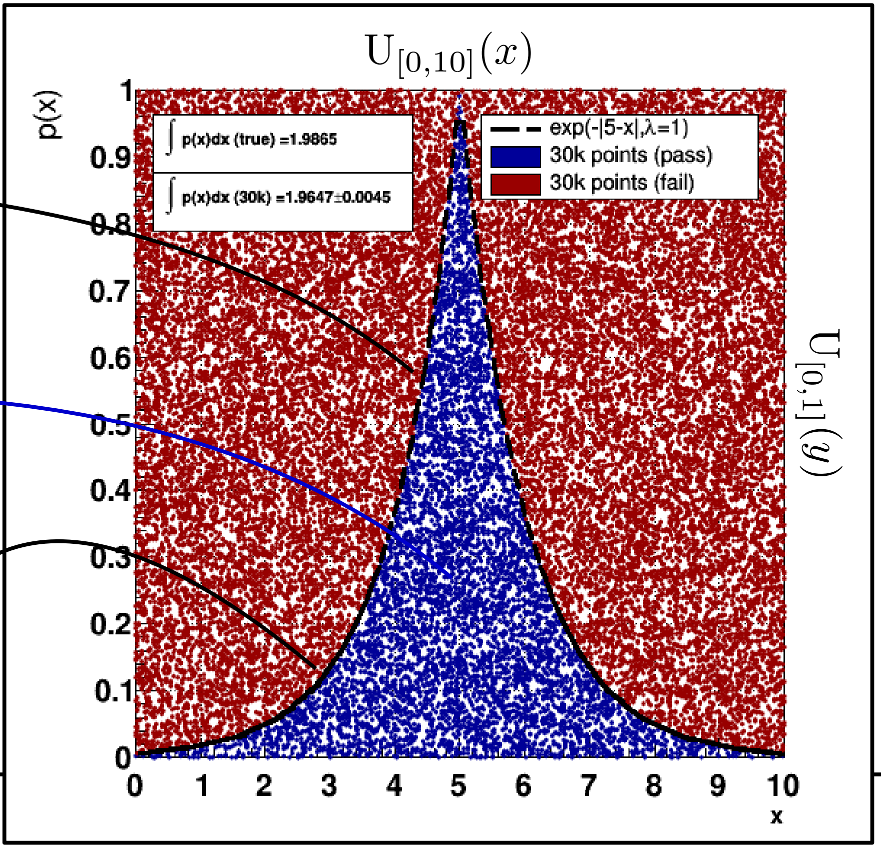
Ein lauffähiges RooT macro finden Sie hier.

- Beispiel Funktion: $p(x) = \exp(-|5-x|)$

Ursprüngliche Wahrscheinlich-keitsdichte

Akzeptierte Inte-grationspunkte (30k stochastic sampling).

Abgeleitetes Histogramm (30k Integrationspunkte)



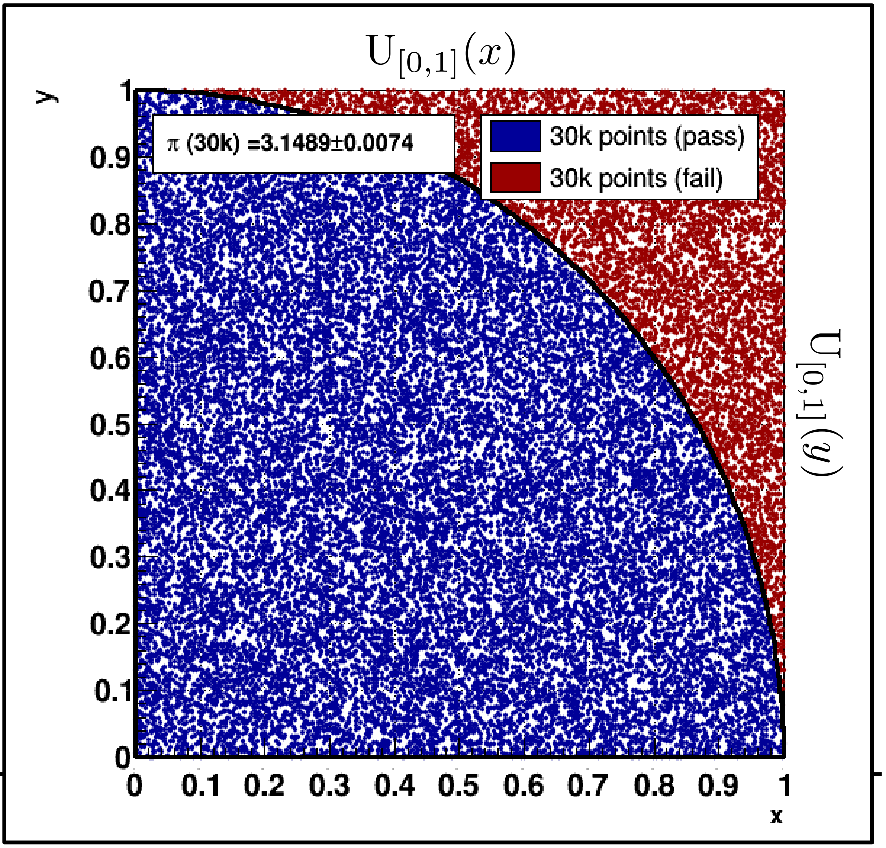
Beispiel – Bestimmung der Zahl π

- Weiteres Beispiel eines einfachen Integrationsproblems ($\rightarrow$ Buffonsches Nadelproblem):

$$\hat{\pi} = 4 \cdot \frac{n_{\mathrm{pass}}}{n_{\mathrm{pass}} + n_{\mathrm{fail}}}$$

(im Beispiel mit 30k stochastic sampling)

Ein lauffähiges RooT macro finden Sie hier.



Priv. Doz. Dr. Roger Wolf
http://ekpwww.physik.uni-karlsruhe.de/~rwolf/

Beispiel – Bestimmung der Zahl π

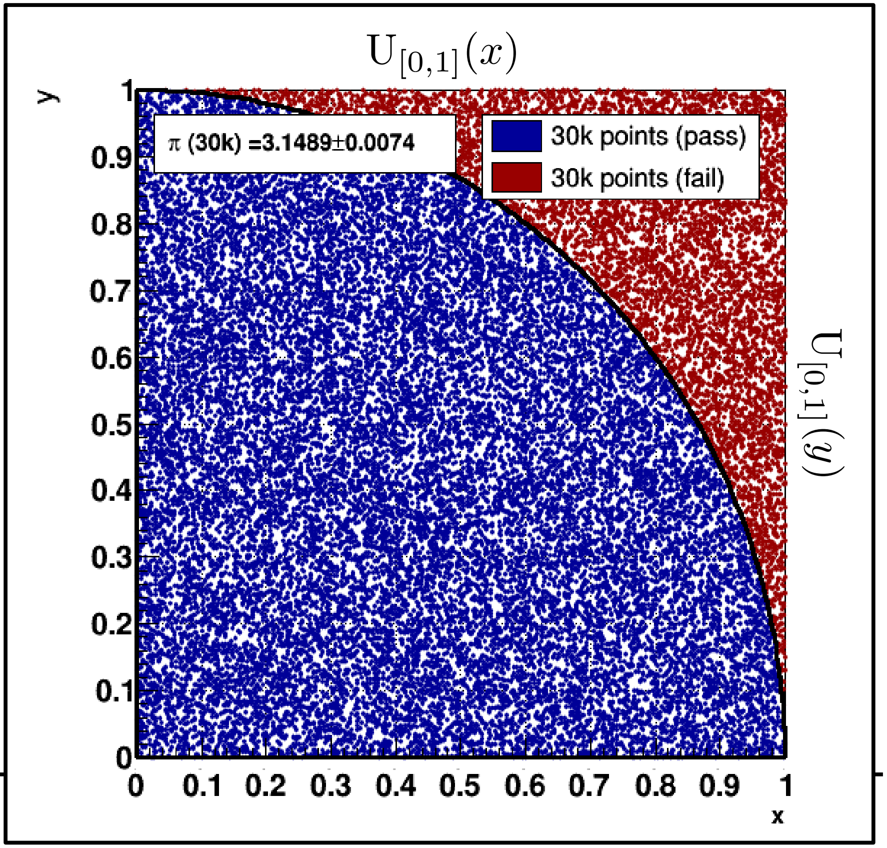
- Weiteres Beispiel eines einfachen Integrationsproblems ($\rightarrow$ Buffonsches Nadelproblem):

$$\hat{\pi} = 4 \cdot \frac{n_{\mathrm{pass}}}{n_{\mathrm{pass}} + n_{\mathrm{fail}}}$$

(im Beispiel mit 30k stochastic sampling)

Georges-Louis Leclerce
de Buffon (1707 – 1788)

Ein lauffähiges RooT macro finden Sie hier.



Beispiel – Bestimmung der Zahl π

- Weiteres Beispiel eines einfachen Integrationsproblems ($\rightarrow$ Buffonsches Nadelproblem):

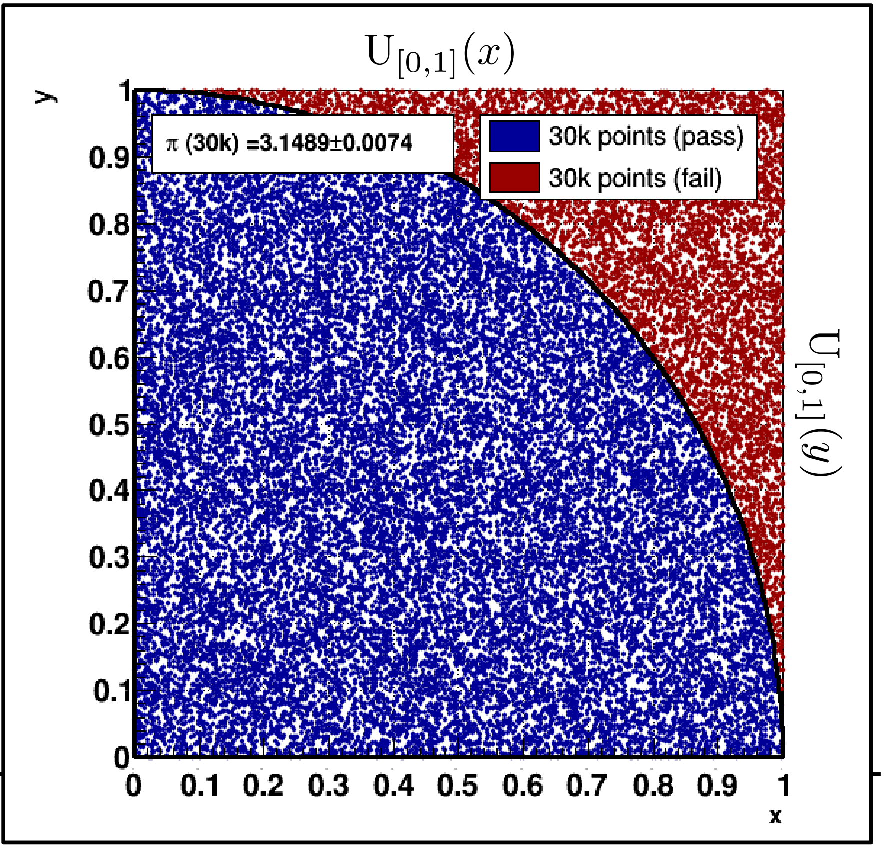
$$\hat{\pi} = 4 \cdot \frac{n_{\text{pass}}}{n_{\text{pass}} + n_{\text{fail}}}$$

(im Beispiel mit 30k stochastic sampling)

Georges-Louis Leclerce
de Buffon (1707 – 1788)

Nicht zu verwechseln mit Gigi Buffon (italienischer Nationaltorhüter)...

Ein lauffähiges RooT macro finden Sie hier.



Straight random sampling

- Ein systematischeres Integrationsverfahren besteht im *straight random sampling*:

$$I = \int\limits_{x_1}^{x_2} \phi(x)\,\mathrm{d}x = ?$$

  - Zur numerischen Integration evaluiere $\phi(x)$ an N gleichverteilten Stützstellen $\{x_i\}$. Damit gilt:

$$\lim_{N\to\infty}\left(\frac{1}{N}\sum_{i=1}^{N}\phi(x_i)\right) = E[x] = \int\limits_{x_1}^{x_2}\phi(x)\,\mathrm{U}_{[x_1,x_2]}\,\mathrm{d}x = \frac{1}{x_2-x_1}\int\limits_{x_1}^{x_2}\phi(x)\,\mathrm{d}x\,;$$

$$I = \int\limits_{x_1}^{x_2}\phi(x)\,\mathrm{d}x = \lim_{N\to\infty}\left(\frac{x_2-x_1}{N}\sum_{i=1}^{N}\phi(x_i)\right)$$

Straight random sampling

- Ein systematischeres Integrationsverfahren besteht im *straight random sampling*:

$$I = \int\limits_{x_1}^{x_2} \phi(x)\,\mathrm{d}x = ?$$

  - Zur numerischen Integration evaluiere $\phi(x)$ an N gleichverteilten Stützstellen $\{x_i\}$. Damit gilt:

$$\lim_{N\to\infty}\left(\frac{1}{N}\sum_{i=1}^{N}\phi(x_i)\right) = E[x] = \int\limits_{x_1}^{x_2}\phi(x)\,\mathrm{U}_{[x_1,x_2]}\,\mathrm{d}x = \frac{1}{x_2-x_1}\int\limits_{x_1}^{x_2}\phi(x)\,\mathrm{d}x\,;$$

$$I = \int\limits_{x_1}^{x_2}\phi(x)\,\mathrm{d}x = \lim_{N\to\infty}\left(\frac{x_2-x_1}{N}\sum_{i=1}^{N}\phi(x_i)\right)$$

  - Die statistische Unsicherheit dieses Verfahrens läßt sich aus der **Varianz** bestimmen:

$$\mathrm{var}[I(N)] = \mathrm{var}\left[\frac{x_2-x_1}{N}\sum_{i=1}^{N}\phi(x_i)\right] = \left(\frac{x_2-x_1}{N}\right)^2 \mathrm{var}\left[\sum_{i=1}^{N}\phi(x_i)\right] = \frac{(x_2-x_1)^2}{N}\mathrm{var}[\phi]$$

Straight random sampling

- Ein systematischeres Integrationsverfahren besteht im *straight random sampling*:

$$I = \int_{x_1}^{x_2} \phi(x)\,\mathrm{d}x = ?$$

- Zur numerischen Integration evaluiere $\phi(x)$ an N gleichverteilten Stützstellen $\{x_i\}$. Damit gilt:

$$\lim_{N\to\infty}\left(\frac{1}{N}\sum_{i=1}^{N}\phi(x_i)\right) = E[x] = \int_{x_1}^{x_2}\phi(x)\,\mathrm{U}_{[x_1,x_2]}\,\mathrm{d}x = \frac{1}{x_2-x_1}\int_{x_1}^{x_2}\phi(x)\,\mathrm{d}x\,;$$

$$I = \int_{x_1}^{x_2}\phi(x)\,\mathrm{d}x = \lim_{N\to\infty}\left(\frac{x_2-x_1}{N}\sum_{i=1}^{N}\phi(x_i)\right)$$

- Die statistische Unsicherheit dieses Verfahrens läßt sich aus der **Varianz** bestimmen:

$$\mathrm{var}[I(N)] = \mathrm{var}\left[\frac{x_2-x_1}{N}\sum_{i=1}^{N}\phi(x_i)\right] = \left(\frac{x_2-x_1}{N}\right)^2 \mathrm{var}\left[\sum_{i=1}^{N}\phi(x_i)\right] = \frac{(x_2-x_1)^2}{N}\mathrm{var}[\phi]$$

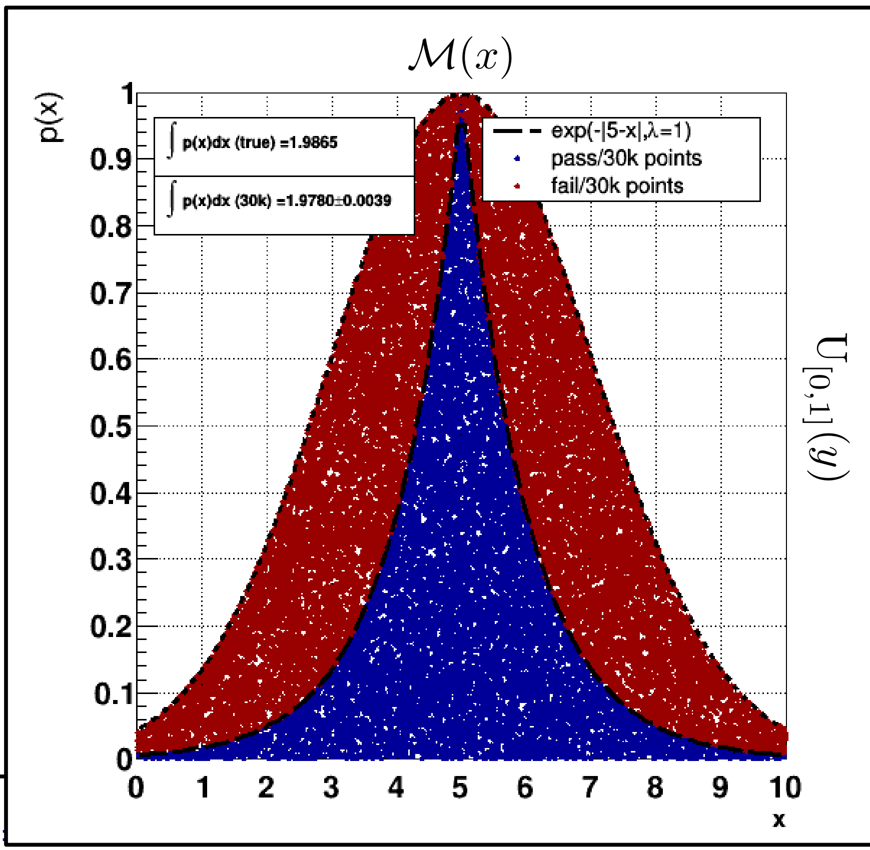
d.h. die Unsicherheit auf $I(N)$ skaliert mit $1/\sqrt{N}$ unabhängig von der Dimension des Integrals.

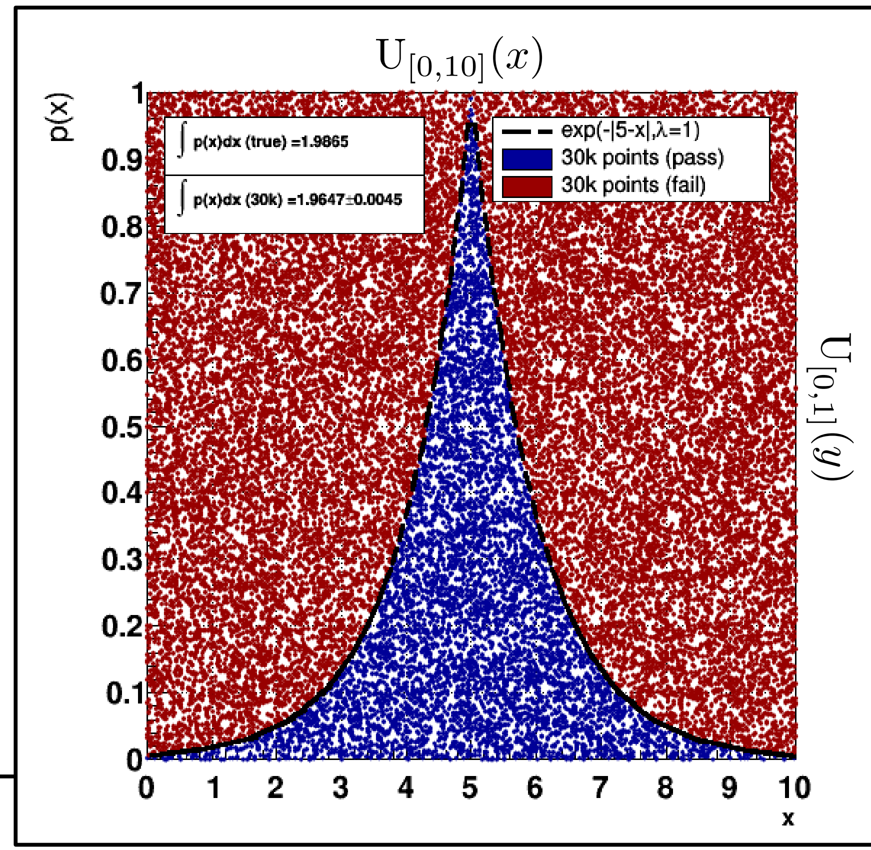
Majoranten Methode

- Die Varianz der Methode relativ zur wahren Wahrscheinlichkeitsdichte/dem wahren Integralwert lässt sich durch Reduktion der *fail* Ereignisse verbessern:

- Beispiel: Wähle $\mathcal{M}(x) = 5\cdot\exp(-\left({(x-5)}/{2}\right)^2)$ als Einhüllende (Majorante). Sample aus Majorante z.B. mit Hilfe der Transformationsmethode.

Ein lauffähiges RooT macro finden Sie hier.

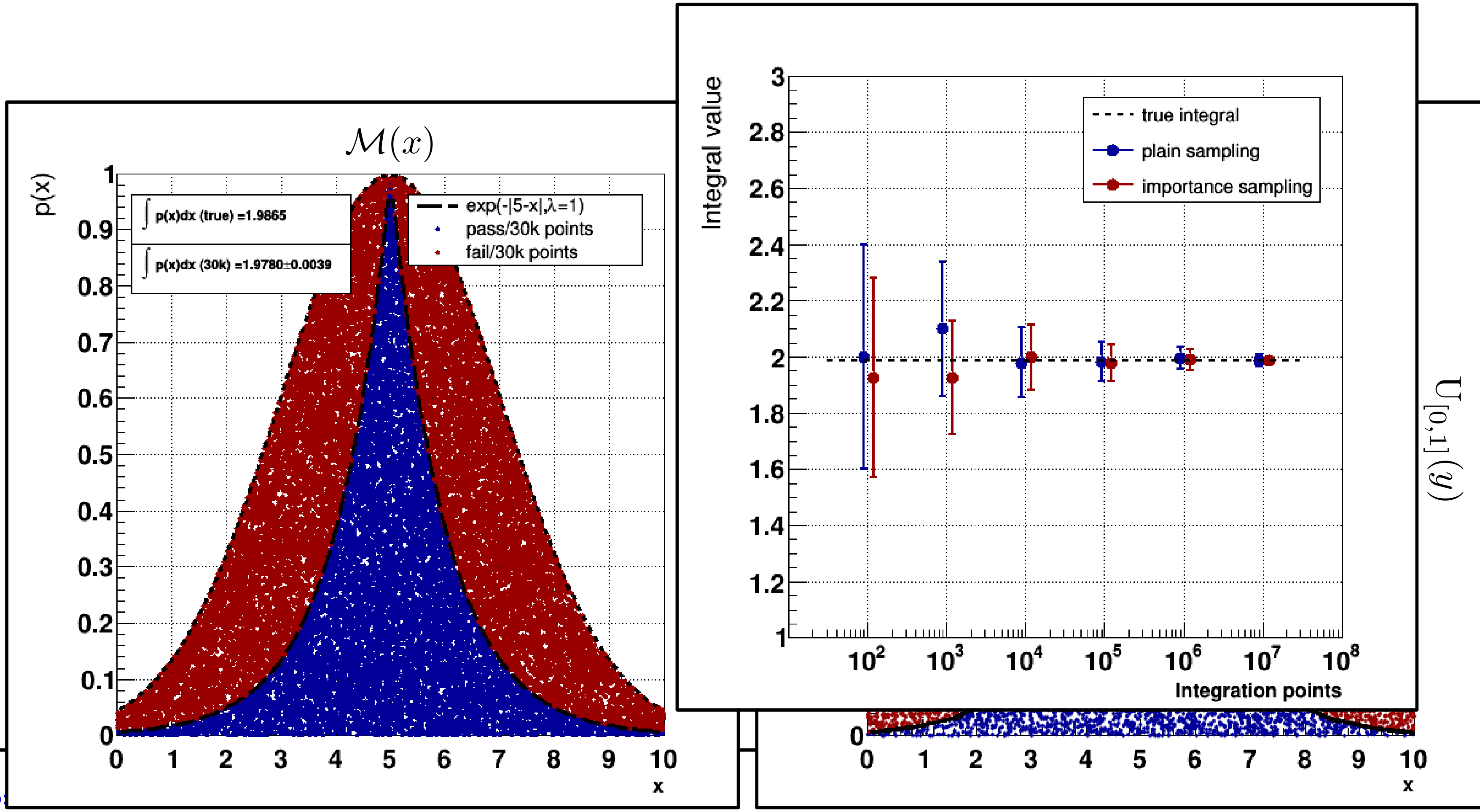




Majoranten Methode

- Die Varianz der Methode relativ zur wahren Wahrscheinlichkeitsdichte/dem wahren Integralwert lässt sich durch Reduktion der *fail* Ereignisse verbessern:

- Beispiel: Wähle $\mathcal{M}(x) = 5 \cdot \exp(-\,(^{(x-5)}/_{2})^2)$ als Einhüllende (Majorante). Sample aus Majorante z.B. mit Hilfe der Transformationsmethode.



Importance sampling

- Die Modifikation des *straight random sampling* sieht wie folgt aus:

$$\lim_{N\to\infty}\left(\frac{1}{N}\sum_{i=1}^{N}\frac{\phi(x_i)}{\mathcal{M}(x_i)}\right) = E_{\mathcal{M}}[x] = \frac{\int\limits_{x_1}^{x_2}\frac{\phi(x)}{\mathcal{M}(x)}\,\mathcal{M}(x)\,\mathrm{d}x}{\int\limits_{x_1}^{x_2}\mathcal{M}(x)\,\mathrm{d}x}\;;$$

$$I = \int\limits_{x_1}^{x_2}\phi(x)\,\mathrm{d}x = \lim_{N\to\infty}\left(\frac{\int\limits_{x_1}^{x_2}\mathcal{M}(x)\,\mathrm{d}x}{N}\sum_{i=1}^{N}\frac{\phi(x_i)}{\mathcal{M}(x_i)}\right)$$

  - Das Konvergenzverhalten als Funktion von N bleibt gleich.
  - Das Integral $\int\limits_{x_1}^{x_2}\mathcal{M}(x)\,\mathrm{d}x$ ist exakt bekannt.
  - Der Mittelwert $\frac{1}{N}\sum_{i=1}^{N}\frac{\phi(x_i)}{\mathcal{M}(x_i)}$ geht gegen 1 und I gegen den bekannten analytischen Wert.

Importance sampling

- Die Modifikation des *straight random sampling* sieht wie folgt aus:

$$\lim_{N\to\infty}\left(\frac{1}{N}\sum_{i=1}^{N}\frac{\phi(x_i)}{\mathcal{M}(x_i)}\right) = E_{\mathcal{M}}[x] = \frac{\int\limits_{x_1}^{x_2}\frac{\phi(x)}{\mathcal{M}(x)}\,\mathcal{M}(x)\,\mathrm{d}x}{\int\limits_{x_1}^{x_2}\mathcal{M}(x)\,\mathrm{d}x}\,;$$

$$I = \int\limits_{x_1}^{x_2}\phi(x)\,\mathrm{d}x = \lim_{N\to\infty}\left(\frac{\int\limits_{x_1}^{x_2}\mathcal{M}(x)\,\mathrm{d}x}{N}\sum_{i=1}^{N}\frac{\phi(x_i)}{\mathcal{M}(x_i)}\right)$$

- Das Konvergenzverhalten als Funktion von N bleibt gleich.
- Das Integral $\int\limits_{x_1}^{x_2}\mathcal{M}(x)\,\mathrm{d}x$ ist exakt bekannt.
- Der Mittelwert $\frac{1}{N}\sum_{i=1}^{N}\frac{\phi(x_i)}{\mathcal{M}(x_i)}$ geht gegen 1 und I gegen den bekannten analytischen Wert.

Jedes weitere Verfahren zur Varianzreduktion besteht entweder darin die numerische Integration durch bekannte analytische Teile zu ersetzen oder das *sampling* auf die Teile von $\phi(x)$ zu konzentrieren, die sich am meisten von 0 unterscheiden.

Vergleich: MC Integration vs. Integration durch Stützstellen

- Die MC Methode bezieht ihre Relevanz aus ihrer Einfachheit & ihrem Konvergenzverhalten.
- Die Konvergenz von Quadraturformeln zur Integration verläuft mit $^1/_{n^{2/d}}$ (wobei n der Anzahl der Stützstellen und d der Dimension des zu integrienden Phasenraums entspricht).
- D.h. ab $d=5$ ist die Monte Carlo Integration Quadraturformeln ihrem Konvergenzverhalten nach überlegen.

2.3 Bootstrap und Ensemble Tests

Während man z.B. im 18. und 19. Jhrd mangels besserer Möglichkeiten Annahmen über Wahrscheinlichkeitsdichten machen musste um (z.T. nur näherungsweise) analystische Rechnungen durchführen zu können, sind im Zeitalter der computergestützen Datenauswertung Ensemble Tests ein wesentlicher Bestandteil zur Auswertung und Einschätzung statistischer Kennzahlen und damit verbundener (physikalischer) Messungen.

Bootstrap-Verfahren

- Bradley Efron, *Bootstrap methods: Another look at the jackknife*, Annals of Statistics, 1979.

- Der Begriff *bootstrap* gründet auf eine Anspielung auf Baron Münchhausen, der sich am eigenen Schopf aus dem Sumpf zog („*to pull oneself up by one‘s bootstrap*“).

- Ursprüngliche Anwendung: Basierend auf wiederholten **Bootstrap-Stichproben** (mit Zurücklegen) wird aus einer einzelnen beobachteten Stichprobe auf die Wahrschein-lichkeitsverteilung (oder Kennzahlen derselben) der Grundgesamtheit geschlossen.

- Die erwartete Verteilungsfunktion strebt dabei der wahren Verteilungsfunktion zu (→ vergleiche VL-01 Folien 47ff).

- **NB:** Diese Methode wird z.B. auch zum training neuronaler Netzwerke nach dem mini batch-Verfahren verwendet.

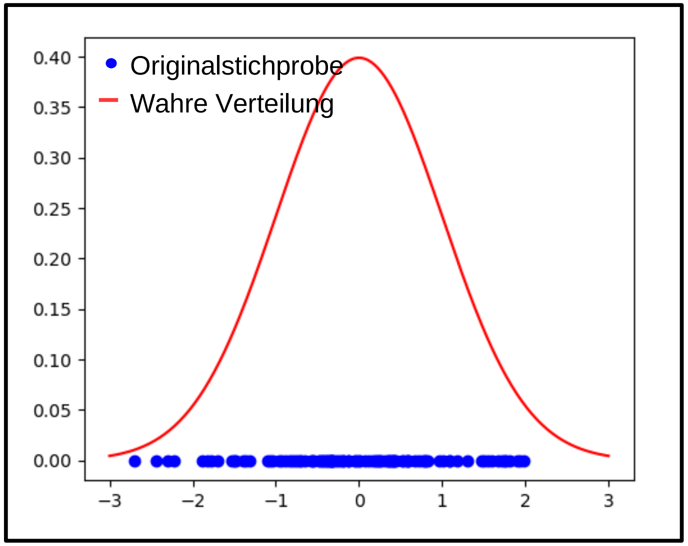
Bootstrap-Verfahren – Beispiel –

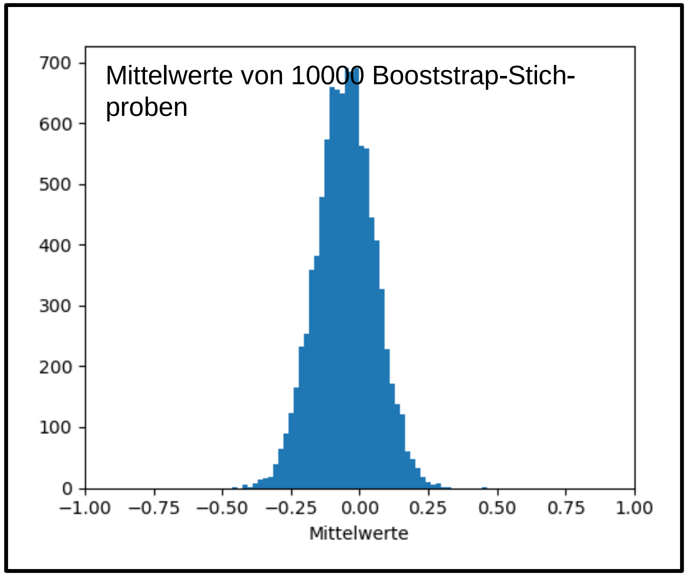
- Wahre Verteilung: $\varphi(x, \mu = 0, \sigma = 1)$

- Originalstichprobe der Länge $N = 100$ (*):
  $\overline{x} = -0.052\,; \quad \sqrt{s^2} = 1.06$

- Erwartungswert und Varianz für $\overline{x}$ aus 10000 Bootstrap-Stichproben aus (*):
  $$\langle\overline{x}\rangle = -0.051\,; \quad \sqrt{\mathrm{var}\left[\langle\overline{x}\rangle\right]} = 0.11$$

- Zum Vergleich: Erwartungswert und Varianz des Mittelwertes der Stichprobe aus wahrer Verteilung:
  $$\langle\overline{x}\rangle = 0 \quad ; \quad \sqrt{\mathrm{var}\left[\langle\overline{x}\rangle\right]} = \frac{\sigma}{\sqrt{100}} = 0.10$$





Bootstrap-Verfahren – Beispiel –

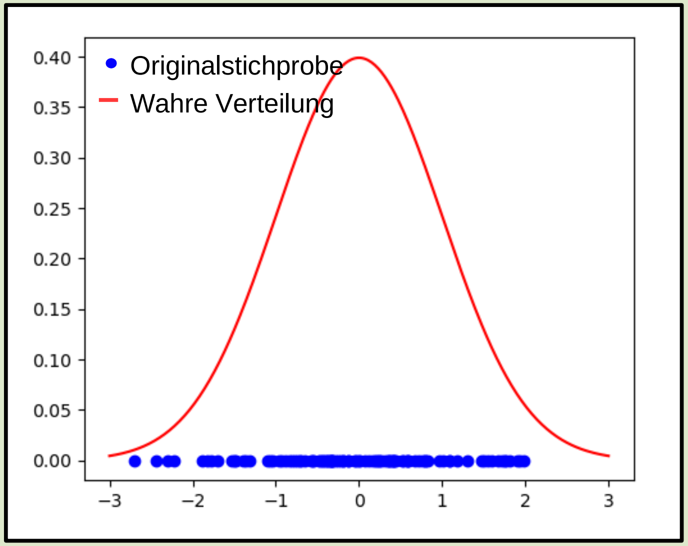
- Wahre Verteilung: $\varphi(x, \mu = 0, \sigma = 1)$

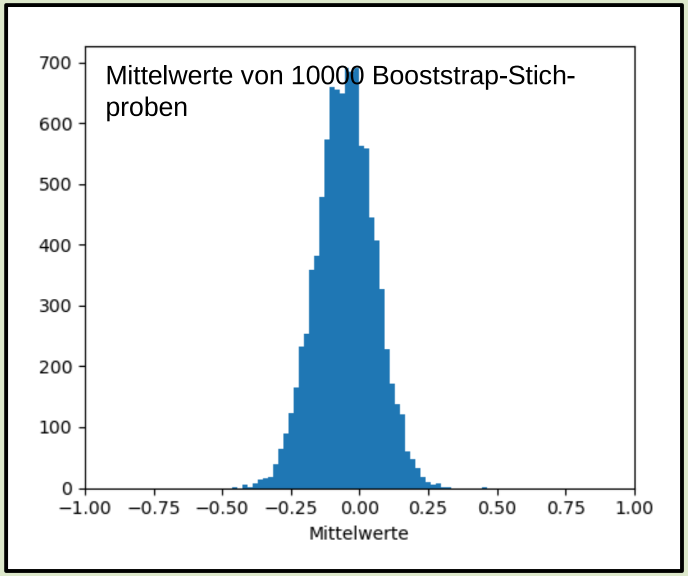
- Originalstichprobe der Länge $N = 100$ (*):
  $\overline{x} = -0.052\,; \quad \sqrt{s^2} = 1.06$

- Erwartungswert und Varianz für $\overline{x}$ aus 10000 Bootstrap-Stichproben aus (*):

$$\langle\overline{x}\rangle = -0.051\,; \quad \sqrt{\mathrm{var}\left[\langle\overline{x}\rangle\right]} = 0.11$$

D.h. 10000 Bootstrap-Stichproben der Länge $N = 100$ aus der originalen Stichprobe. Die Bootstrap-Stichprobe erfolgt mit Zurückle-gen. Es kommt also nicht immer der geiche Mittelwert dabei heraus.





- Zum Vergleich: Erwartungswert und Varianz des Mittelwertes der Stichprobe aus wahrer Verteilung:

$$\langle\overline{x}\rangle = 0 \quad ; \quad \sqrt{\mathrm{var}\left[\langle\overline{x}\rangle\right]} = \frac{\sigma}{\sqrt{100}} = 0.10$$

Ensemble Tests

- Oft ist die Messung eines Parameters Ergebnis einer einzelnen Messung. Frage: ist das Ergebnis repräsentativ? Sind die Unsicherheiten korrekt abgeschätzt – eventuelle Korrelationen bei der Fehlerrechnung berücksichtigt?

- Eine konzeptionell klare Antwort erhalten Sie durch die Durchführung eines **Ensemble Tests**. Dieser geht wie folgt vonstatten:

  - Setze eine hypothetische Wahrscheinlichkeitsdichte als gegeben voraus.

  - Simuliere das Experiment so oft wie möglich ($\rightarrow$ Pseudo-Experimente).

  - Analysiere jedes Pseudo-Experiment genau so wie das durchgeführte Experiment.

  - Extrahiere die statistischen Eigenschaften aus diesem Ensemble an Pseudo-Experimenten.

Ensemble Tests – Beispiel 1 –

- **Experiment:**
  - Radioaktiver Zerfall eines Präparats.

$$N(t) = N_0\, e^{-\lambda\, t}\,; \quad \text{bekannt: } N_0\,, \lambda = 0.2\,\mathrm{s}^{-1}$$

$$\dot{N}(t) = -\lambda N(t)$$

    - 100 Messpunkte,
    - Histogrammiert mit vorgegebener Anzahl an Bins,
    - Bestimme λ aus Anpassung an die Messpunkte.

- **Ensemble Test:**
  - Würfle z.B. 1000 Pseudo-Experimente mit vorgegebener Wahrscheinlichkeitsdichte.
  - Histogrammiere den Ausgang jedes Pseudo-Experiments wie für die durchgeführte Messung und passe die angenommene (Exponential-)Verteilung an.
  - Speichere den ermittelten Wert von λ ab und bestimme aus dem Ensemble Test Mittelwert und Varianz des Mittelwerts.

Ensemble Tests – Beispiel 1 –

- **Ergebnis des Ensemble Tests:**

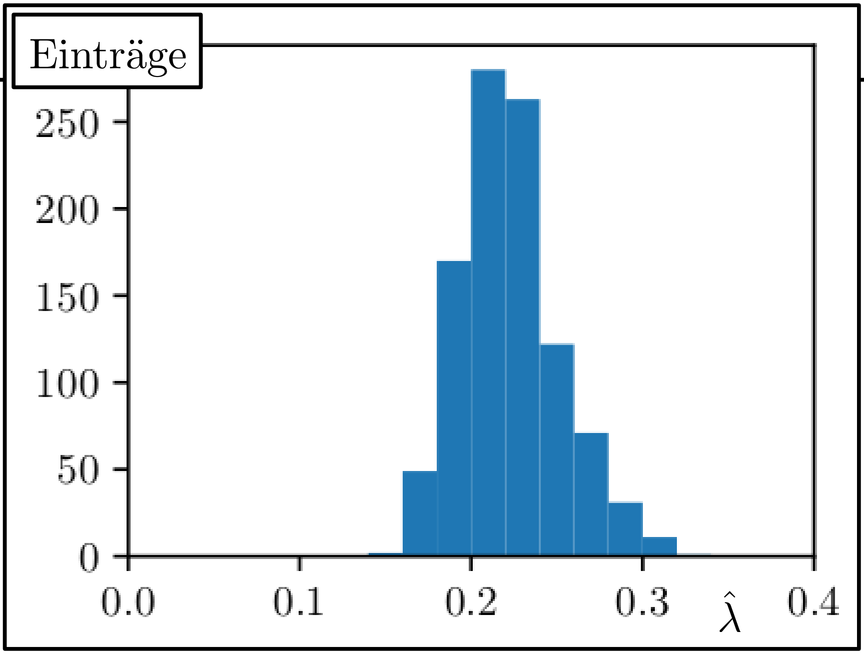
  $\langle\hat{\lambda}\rangle = 0.223 \pm 0.029\,\mathrm{s}^{-1}$

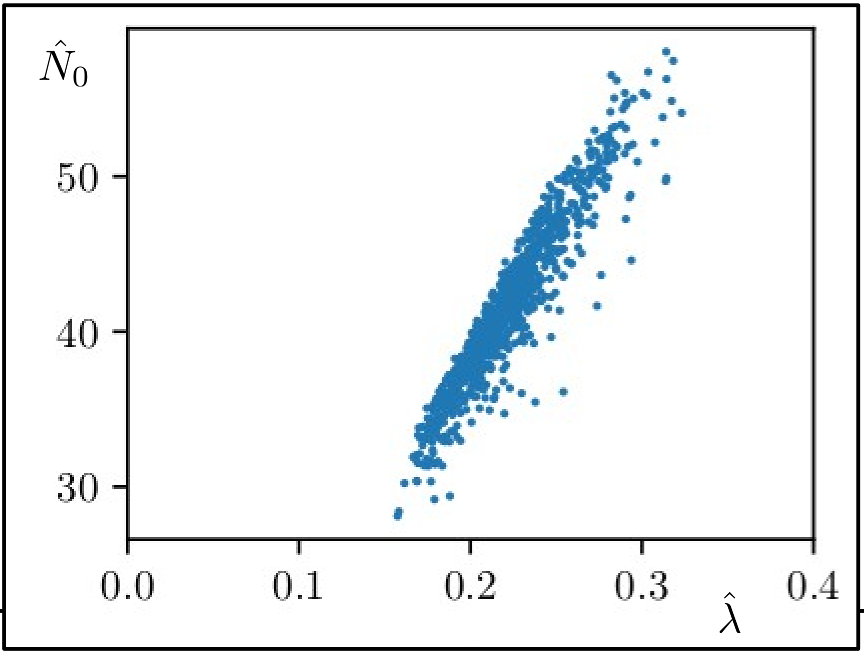
  bei vorgegebenem wahrem Wert von:

  $\lambda = 0.2$

  - Die Werte von $\hat{\lambda}$ und $\hat{N}_0$ sind korreliert und die Verteilung von $\hat{\lambda}$ nicht normal-verteilt.

  - Mit Hilfe einer analytischen Rechnung hätten Sie bei diesem einfachen Beispiel bereits Schwierigkeiten bei der Interpretation der Unsicherheiten.





Ensemble Tests – Beispiel 2 –

- Physikalische Großexperimente (z.B. am LHC) sind zu komplex, um analytische Vorhersagen zuzulassen.

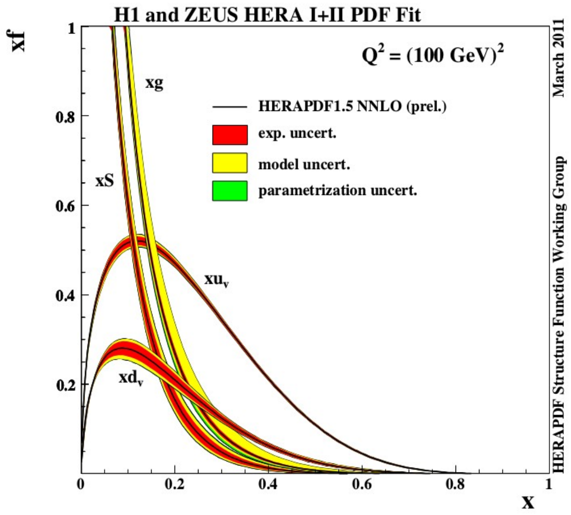
- Insbesondere Experimente der Teilchenphysik eignen sich jedoch hervorragend für die Simulation mit Hilfe der Monte Carlo Methode.

  - So gut wie alle Elemente der Simulation sind stochastischer Natur.

  - Nicht nur die statistische Gesamtheit sondern jedes einzelne Ergebnis („Ereignis“) der Simulation hat die Bedeutung eines möglichen Ausgangs des Experiments.

Monte Carlo Methoden in der Teilchenphysik

Wahrscheinlichkeit ein bestimmtes Parton im Proton anzutreffen;



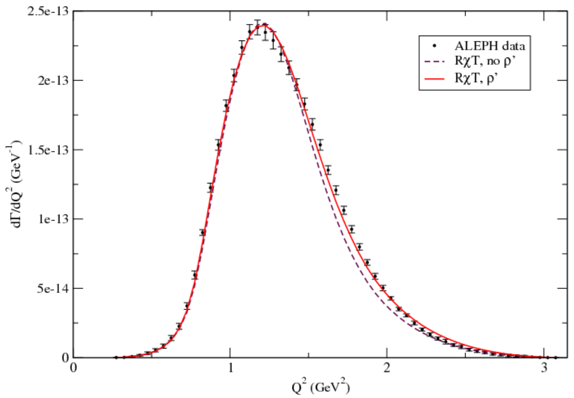
Monte Carlo Methoden in der Teilchenphysik

Wahrscheinlichkeit einen
bestimmten Endzustand
zu beobachten:

$$\mathcal{M}_{ij} = \frac{\mathcal{G}^2}{2} \mathcal{L}^{\mu\nu}_{\tau\tau} \mathcal{L}^{ee}_{\mu\nu}$$

Monte Carlo Methoden in der Teilchenphysik

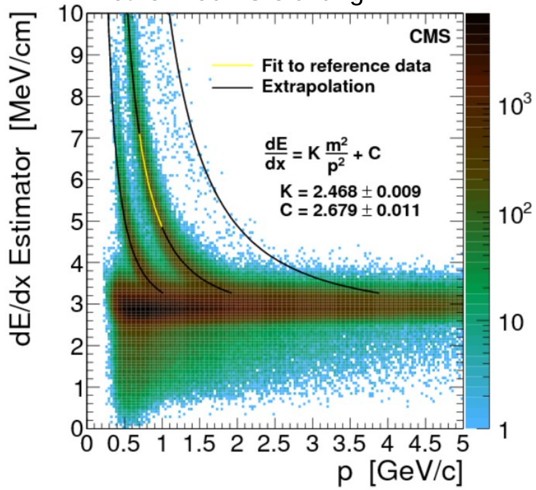
Wahrscheinlichkeit für
den Zerfall $\tau \rightarrow 3\pi\nu_\tau$



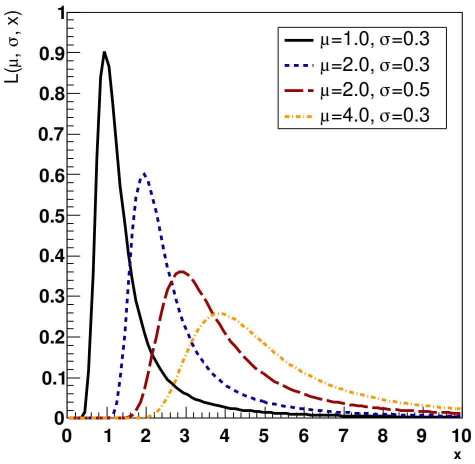
Monte Carlo Methoden in der Teilchenphysik

Energieverlust eines geladenen Teilchens in einem Medium.

Bethe-Bloch Gleichung



Landau Verteilung



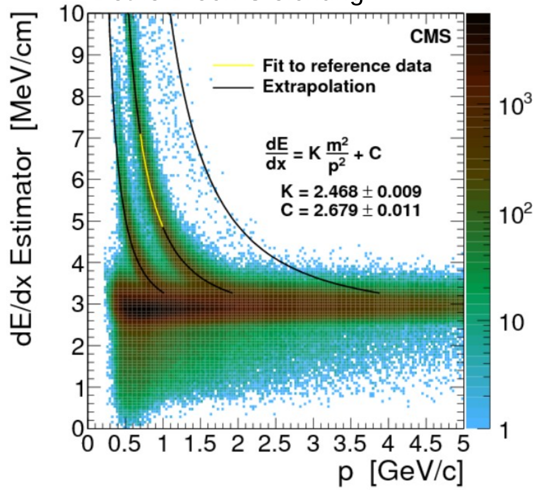
Priv. Doz. Dr. Roger Wolf
http://ekpwww.physik.uni-karlsruhe.de/~rwolf/

Monte Carlo Methoden in der Teilchenphysik

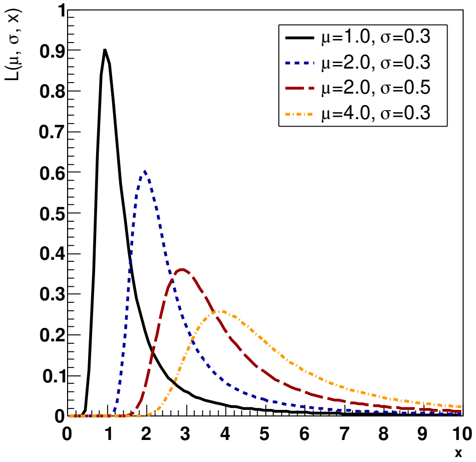
Energieverlust eines geladenen Teilchens in einem Medium.

- Alles stochastisch unabhängige Prozesse
- Jeder akzeptierte Integrationspunkt könnte ein echter Ausgang des Experiments sein.

Bethe-Bloch Gleichung

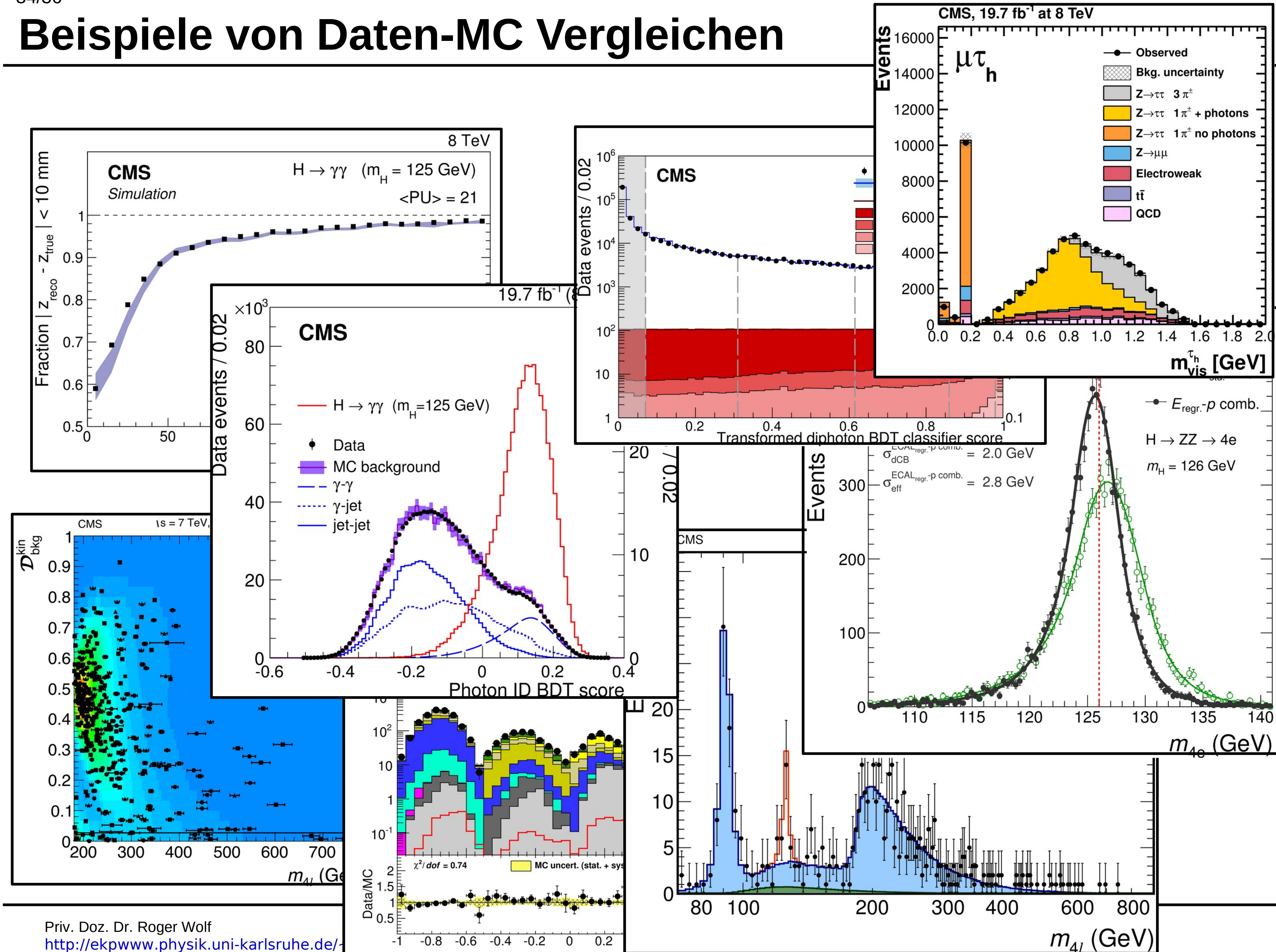


Landau Verteilung



Priv. Doz. Dr. Roger Wolf
http://ekpwww.physik.uni-karlsruhe.de/~rwolf/

Beispiele von Daten-MC Vergleichen



Bedeutung von Ensemble Tests

- Viele analytische Methoden der Statistik setzen bestimmte Wahrscheinlichkeitsdichten (z.B. Normalverteilung oder χ^2-Verteilung) voraus. Diese Voraussetzungen sind nicht immer erfüllt und oft nur schwer nachzuweisen!

- Diese Methoden gründen aus der Zeit als Ensemble Tests sehr aufwändig und daher i.a. nicht möglich waren.

- Vorteil aller *sampling* Methoden (wie dem Bootstrap-Verfahren und Ensemble Tests) ist, dass sie keine solche Annahmen über die Wahrscheinlichkeitsverteilung der Grundgesamtheit machen. Getroffene Annahmen können ohne konzeptionelle Schwierigkeiten systematisch variiert werden.

- Soweit Sie die Möglichkeit dazu haben, sollten Sie einen Ensemble Test einer analytischen Lösung vorziehen oder zumindest die analytische Lösung durch einen Ensemble Test supplementieren.

Zusammenfassung

- Zufallszahlengeneratoren.

- Transformation gleichverteilter Zufallszahlen auf beliebige Wahrscheinlichkeitsdichten und numerische Integration.

- Bootstrap-Verfahren und Ensemble Tests.